



US005082767A

**United States Patent [19]****Hatfield et al.****Patent Number: 5,082,767****Date of Patent: Jan. 21, 1992****[54] CODON PAIR UTILIZATION**

[76] Inventors: **G. Wesley Hatfield**, 718 Goldenrod, Corona Del Mar, Calif. 92625; **George A. Gutman**, 3080 Tyler Way, Costa Mesa, Calif. 92626

[21] Appl. No.: **316,321**

[22] Filed: **Feb. 27, 1989**

[51] Int. Cl.<sup>5</sup> ..... **C12Q 1/68**

[52] U.S. Cl. ..... **435/6; 435/69.1; 435/91; 435/172.3; 436/501; 536/26; 536/27; 536/28; 935/17; 935/78; 935/88**

[58] Field of Search ..... **435/6, 91, 69.1, 172.3; 436/501; 536/26, 27, 28; 935/17, 78, 88**

**[56] References Cited****U.S. PATENT DOCUMENTS**

4,184,917	1/1980	Dorner et al.	.....	435/68
4,190,495	2/1980	Curtiss, III	.....	435/172
4,262,090	4/1981	Colby, Jr. et al.	.....	435/91
4,293,652	10/1981	Cohen	.....	435/172
4,321,365	3/1982	Wu et al.	.....	536/27
4,332,892	6/1982	Ptashne et al.	.....	435/68
4,332,898	6/1982	Reusser	.....	435/172
4,338,397	7/1982	Gilbert et al.	.....	435/68
4,342,832	8/1982	Goeddel et al.	.....	435/172
4,348,477	9/1982	Nakano et al.	.....	435/172
4,349,629	9/1982	Carey et al.	.....	435/172
4,350,764	9/1982	Baxter et al.	.....	435/69
4,351,901	9/1982	Bahl	.....	435/91
4,356,270	10/1982	Itakura	.....	435/317
4,359,535	11/1982	Pieczenick	.....	435/317
4,363,877	12/1982	Goodman et al.	.....	435/317
4,371,625	2/1983	Tiollais	.....	435/317
4,394,443	7/1983	Weissman et al.	.....	435/6
4,396,601	8/1983	Salser et al.	.....	424/94
4,399,216	8/1983	Axel et al.	.....	435/6
4,403,035	9/1983	Anderson et al.	.....	435/172
4,468,464	8/1984	Cohen et al.	.....	435/317

**OTHER PUBLICATIONS**

Shaper, "Constraints on Codon Context in *Escherichia coli* Genes, Their Possible Role in Modulating the Efficiency of Translation", *J. Mol. Biol.* 188:555-564 (1985). Yarus et al., "Sense Codons Are Found in Specific Contexts", *J. Mol. Biol.* 182:529-540 (1985).

Ausubel et al., Current Protocols in Molecular Biology, vol. 1, Chap. 8, (Wiley Interscience, 1988).

Staden, R., *Nucleic Acids Res.* 12:551-567 (1984).

R. Nussinov, "Eukaryotic Dinucleotide Preference Rules and Their Implications for Degenerate Codon Usage", *J. Mol. Biol.* 149:125-131 (1981).

Bossi, L. and J. R. Roth, "The Influence of Codon Context on Genetic Code Translation", *Nature* 286:123-127 (1980).

Taniguchi, T. and C. Weissman, "Inhibition of Q $\beta$  RNA 70S Ribosome Initiation Complex Formation by an Oligonucleotide Complementary to the 3' Terminal Region of *E. coli* 16S Ribosomal RNA", *Nature* 275:770-772 (1978).

Kolaskar, A. S. and B. V. B. Reddy, "Contextual Constraints on Codon Pair Usage: Structural and Biological Implications", *Journal of Biomolecular Structure and Dynamics* 3(4):725-738 (1986).

Hatfield et al., "Codon Pair Utilization in *E. coli* is Non-Random", *Journal of Cellular Biochemistry*, Supplement 13 C, (1989).

**Primary Examiner—Robert A. Wax**

**Assistant Examiner—Ardin H. Marschel**

**Attorney, Agent, or Firm—Knobbe, Martens, Olson & Bear**

**[57] ABSTRACT**

A method for determining the pattern of nonrandom codon pair usage of an organism, comprising the steps of obtaining nucleotide sequence data for the organism, determining from the data the number of codons represented in at least a portion of the sequence and the frequency of usage of at least some codons in the portion, determining from the frequency the expected number of occurrences of at least some codon pairs, if they are paired in a random manner, and comparing the expected number with the actual number of occurrences to determine relative codon pairing preferences. The codon pairings of organisms are highly nonrandom, and differ from organism to organism. This information is used to construct and express altered or synthetic genes having desired levels of translational efficiency, to determine which regions in a genome are protein coding regions, to introduce translational pause sites into heterologous genes, and to ascertain relationship or ancestral origin of nucleotide sequences.

**44 Claims, 4 Drawing Sheets**

